Running Head: GENDERED RATIONALITY

Rationality is Gendered
Supplemental Online Materials

Olivia Pavco-Giaccia, Martha Fitch Little, Jason Stanley, and Yarrow Dunham
Department of Psychology, Yale University, New Haven, Connecticut, USA

Corresponding Author:

yarrow.dunham@yale.edu

Summary

This supplement includes the following:

1. Additional details on stimulus selection, including the list of all word stimuli used in the IAT in Study 1a and 1b and their frequency attributes; additional details on the photographs used in Study 2.
2. Details concerning the Quadruple Process Model analysis of the IAT data in Study 1a.
3. Additional analyses looking at gender differences on explicit measures conducted on the pooled data from Studies 1a, 1b, 2, and 3.
4. Details concerning the ratings of academic disciplines and subsequent Principle Components Analysis used to characterize the academic disciplines in Study 3.

1. Stimuli Selection

Studies 1a and 1b

Studies 1a and 1b used the same verbal stimuli. The category stimuli designed to invoke the categories *male* and *female*, respectively, were taken from past research with the IAT, and were:

Male: Man, Boy, Father, Male, Grandpa, Husband, Son, Uncle

Female: Woman, Girl, Mother, Female, Grandma, Wife, Daughter, Aunt

The attribute stimuli designed to invoke the categories *thinking* and *feeling*, respectively, were produced by the authors based on synonyms of those focal terms and were then roughly matched on valence, frequency, and length, using the word ratings from Warriner, Kuperman, and Brysbaert (2013). The items and the associated statistical tests for each to establish non-difference on each of these dimensions are:

Rational: Rational, Deliberate, Calculate, Think, Logical, Reason

Emotional: Emotional, React, Intuition, Feel, Empathic, Sense

Table S1: Characteristics of Word Stimuli in IATs (Studies 1a and 1b)

Word	Category	Valence	Log Frequency	Length
rational	Thinking	5.85	3.82	8
deliberate	Thinking	4.8	3.63	10
think	Thinking	6.68	5.81	5
reason	Thinking	6.27	4.92	6
calculate	Thinking	4.89	3.50	9
logical	Thinking	6.38	3.80	7
Thinking	**Average**	**5.81**	**4.25**	**7.5**
emotional	Feeling	6.62	4.41	9
react	Feeling	5.15	3.86	5
feel	Feeling	6.27	5.19	4
intuition	Feeling	6.42	3.30	9
sense	Feeling	6.09	5.04	5
empathic	Feeling	7.29	2.66	8

Feeling	**Average**	**6.31**	**4.08**	**6.67**

Study 2

Faces drawn from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) were matched on unusualness, sadness, attractiveness, age, and happiness; they were, however, well-differentiated on femininity and masculinity. Table S2 provides a list of items and their ratings on each of these dimensions.

Table S2: Mean Stimulus Ratings for Photographs used in Study 2 (AMP)

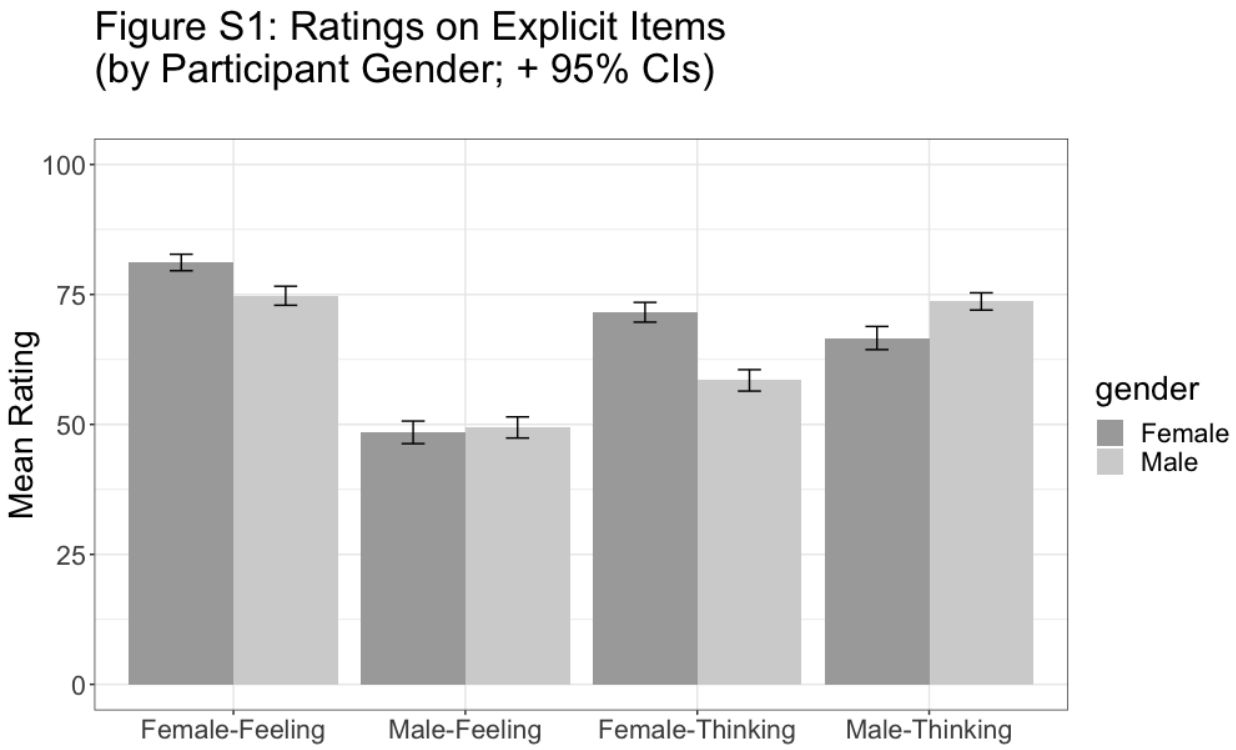
CFD Stimulus Code	Gender	Unusualness	Attractiveness	Age	Sadness	Happiness	Femininity	Masculinity
WF-039	F	2.03	3.35	23.12	3.12	2.48	4.55	2.14
WF-003	F	2.63	5.02	26.09	2.58	3.55	5.56	2.00
WF-011	F	1.78	4.06	23.75	2.52	2.46	5.15	1.75
WF-006	F	2.82	3.46	24.03	3.75	2.24	4.67	2.06
WF-038	F	2.18	3.27	28.27	2.00	2.88	4.14	2.53
WF-027	F	2.05	4.63	21.58	2.08	3.05	5.23	1.58
Female	**Average**	**2.25**	**3.97**	**24.47**	**2.67**	**2.78**	**4.88**	**2.01**
WM-003	M	2.79	3.57	22.98	3.05	2.76	2.02	4.91
WM-004	M	2.23	4.68	25.40	2.74	2.80	2.06	4.76
WM-006	M	2.86	3.65	25.53	2.54	2.88	2.63	4.74
WM-009	M	2.42	4.18	23.76	2.65	2.26	2.21	4.64
WM-029	M	1.98	4.59	28.63	2.56	2.63	1.73	5.24
WM-024	M	2.04	3.75	20.67	2.81	2.86	2.26	4.16
Male	**Average**	**2.39**	**4.07**	**24.49**	**2.73**	**2.70**	**2.15**	**4.74**

2. Quad Model Analysis:

Due to the relative nature of the IAT standard scoring criteria based on differences in reaction time are not able to dissociate the two contributing associations (in this case the proposed link between *male* and *rationality* and *emotionality* and *female.* However, a recent data-analytic procedure seeks to remedy this gap: the *Quadruple Process Model*, or Quad Model (Conrey, Sherman, Gawronski, Hugenberg, & Groom, 2005). The Quad Model analyzes error rates rather than reaction times, building a multinomial processing tree model that estimates an independent parameter estimate for the influence of each association. A Quad Model analysis of our data suggested that our results reflected the joint influence of both associations, as the estimates of each parameter (and their 95% confidence intervals) both exceeded 0: $A_{\text{male-rational}}$ = .10, 95% CI [.083; .13]; $A_{\text{female-emotional}}$ = .071, 95% CI [.050; .093]. Unfortunately, the goodness-of-fit test assessing whether the model fit the data suggested poor fit, G^2 = 207.14, df = 3, p < .001. It has been argued that this test is overly conservative with larger data sets such that it rejects the null of satisfactory fit too frequently (Gonsalkorale, Sherman, & Klauer, 2009), and on that reading these results could stand. However, given the uncertainty surrounding this claim we elected to conduct a second study (Study 1b) to examine this issue more directly. The fact that Study 1b and the Quad Model analysis provide convergent evidence is therefore reassuring.

3. Combined analyses of explicit ratings focusing on gender differences

Participants in all four data collections (Studies 1a and 1b, 2, and 3) completed an identical set of explicit rating items. To gain insight into gender differences in these items we pooled the sample across the three data collections, producing a final sample of 946 participants (490 male, 451 female, 4 unreported). Mean ratings on each of the four items are presented in Figure S1. This aggregated analysis revealed consistent gender differences on three of the four items.



More specifically, while there was no difference between men and women in terms of the tendency to associate *male* with *feeling*, $t(928.39) = .62$, $p = .53$, there were moderate tendencies for women to more strongly endorse a link between *female* and *feeling* ($t(929.59) = 5.10$, $p < .0001$, $d = .34$) and for men to more strongly endorse a link between *male* and *thinking* ($t(848.33) = -4.98$, $p < .0001$, $d = -.32$). Further, there was a strong tendency for women to more strongly endorse a link between *female* and *thinking* than did men, $t(937.56) = 9.18$, $p < .0001$, $d = .60$.

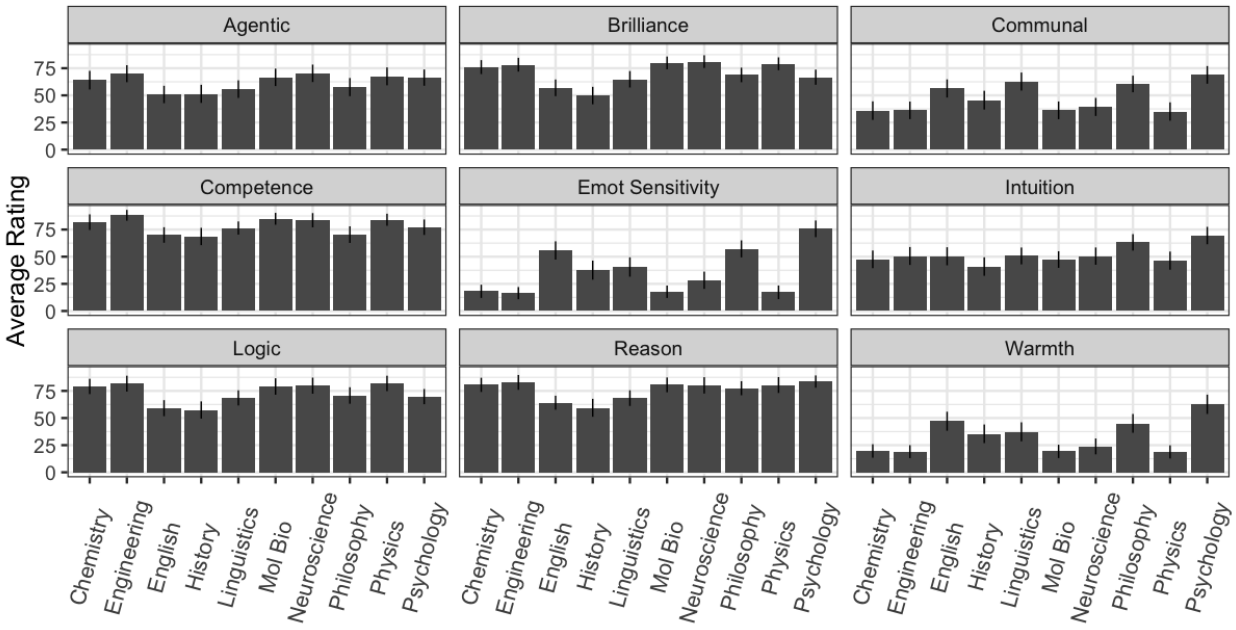
Another way to consider these differences is to note that women saw a large difference in terms of whether *male* versus *female* as associated with *feeling* but saw very little difference in the extent to which *male* versus *female* are associated with *thinking*; by contrast, men showed very strong tendencies to see both *thinking* and *feeling* as heavily gendered.

4. Details of discipline ratings and PCA in Study 3

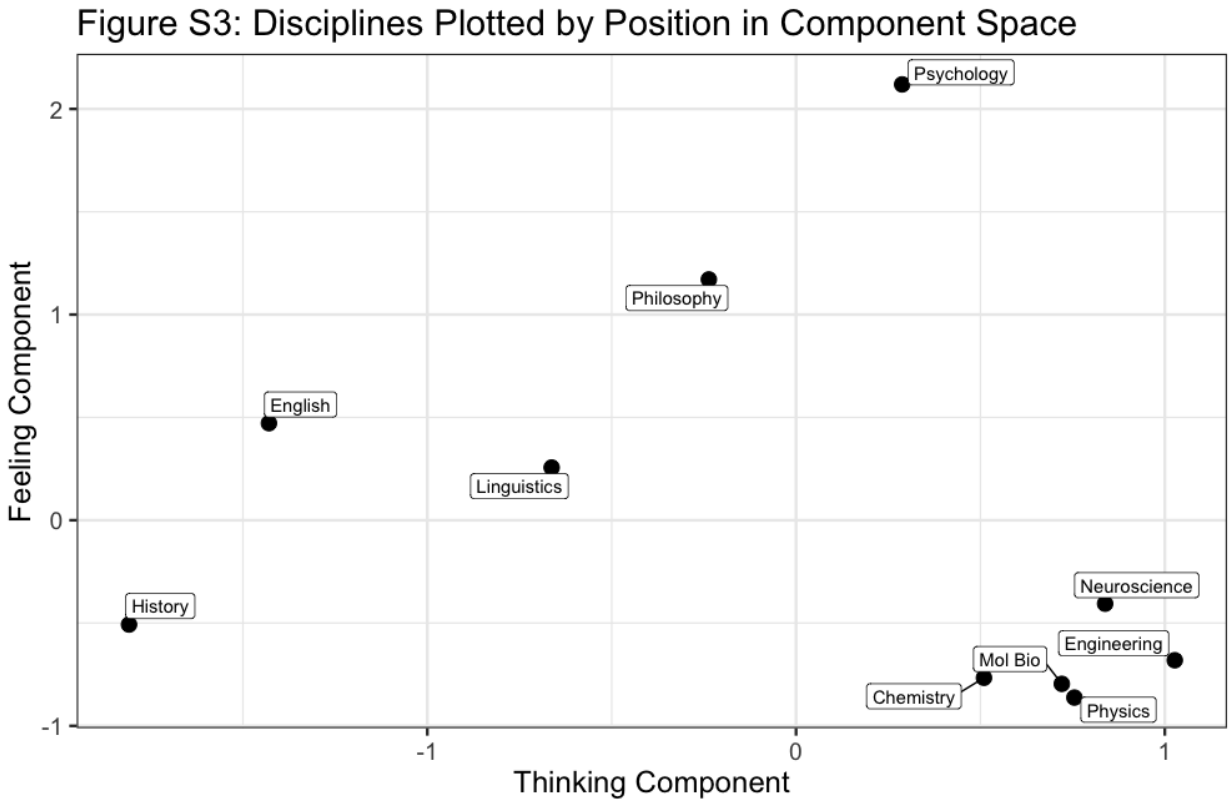
Field ratings.

We secured ratings of the 10 academic disciplines from participants on Amazon's Mechanical Turk. 50 participants were recruited, of whom 2 exited the survey before completing any ratings. One further participant was excluded for responding with a value of 100 (the scale maximum) to every question, leaving 47 usable participants. Each of the 10 disciplines was rated by each participant on the extent to which each of 9 dimensions was required for success in that field (agency, competence, logic, reason, brilliance, communal, emotional sensitivity, intuition, and warmth). Average ratings for each field on each dimension are presented in Figure S2.

Figure S2: Average Rating (and 95% CI) for each Discipline on each Trait



These ratings were averaged and then submitted to a Principle Component Analysis. Inspection of the Scree Plot and eigenvalues strongly suggested a two-factor solution, with the first two factors accounting for 96% of ratings variance and the third potential factor adding only 2% of additional variance. We therefore fit a 2-component solution using an oblimin rotation to allow the two dimensions to be correlated; indeed, they were negatively correlated at r = -.25. Figure S3 provides a plot of the 10 disciplines in this 2-dimensional component space.



Identifying Top 20 Departments

Top 20 departments were chosen by combining the lists of *US News and World Report* and the *QS Top Univerities* rankings to identify the 20 schools in common if both rankings were present or using the available rankings when only one of these sources listed a given discipline. While limited in many respects, we emphasize that this procedure was conducted in a data-blind manner before any primary analyses were conducted and at the very least serves as a rough sense of schools ranked quite highly in each of the relevant disciplines.

References in this Supplement

Conrey, F. R., Sherman, J. W., Gawronski, B., Hugenberg, K., & Groom, C. J. (2005). Separating Multiple Processes in Implicit Social Cognition: The Quad Model of Implicit Task Performance. *Journal of Personality and Social Psychology*, *89*(4), 469–487. http://doi.org/10.1037/0022-3514.89.4.469

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, *45*(2), 1–5. http://doi.org/10.1016/j.jesp.2008.11.004

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. http://doi.org/10.3758/s13428-014-0532-5

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x